

(19)日本国特許庁 (J P)

(12) 実用新案登録公報 (Y 2) (11)実用新案登録番号

第2500761号

(45)発行日 平成 8 年(1996) 6 月12日

(24)登録日 平成 8 年(1996) 3 月28日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/02	3 0 1		G 1 0 L 3/02	3 0 1 B

(全 4 頁)

(21)出願番号 実願平6-3041
特願昭61-122324の変更
(22)出願日 昭和61年(1986) 5 月29日
(65)公開番号 実開平6-82700
(43)公開日 平成 6 年(1994)11月25日

(73)実用新案権者 000170598
株式会社アルファ
神奈川県横浜市金沢区福浦 1 丁目 6 番 8 号
(72)考案者 白井 俊之
東京都大田区蒲田 2 丁目 8 番 2 号 国産
金属工業株式会社内
(74)代理人 弁理士 瀧野 秀雄 (外 1 名)

審査官 渡邊 聡

(56)参考文献 特開 昭57-90750 (J P, A)
特開 昭58-70287 (J P, A)
特開 昭57-179898 (J P, A)
特開 昭58-125099 (J P, A)

(54)【考案の名称】 音声認識装置

1

(57)【実用新案登録請求の範囲】

【請求項 1】 音声入力レベルに応じた大きさの音声変換電圧信号を発生する音声-電気変換手段と、
該音声-電気変換手段からの音声変換電圧信号を増幅する増幅手段と、
該増幅手段の利得を設定する利得設定手段と、
前記増幅手段を介して前記音声-電気変換手段からの音声変換電圧信号を受け取り音声認識を行う音声認識部と、
前記音声-電気変換手段からの音声変換電圧信号を予め定めた基準電圧と比較し、前記音声変換電圧信号が基準電圧以上となったときセット信号を出力する比較手段と、
音声発声者の操作により発生されるリセット信号によりリセットされ、リセット状態にあるとき前記比較手段か

2

らのセット信号の入力により一定時間持続する時間信号を出力するタイマ手段と、
該タイマ手段が時間信号を出力している間に入力する前記音声変換電圧信号のピーク値を検出するピーク検出手段とを備え、
前記利得設定手段が、前記ピーク検出手段により検出したピーク値に反比例した値に前記増幅手段の利得を設定することを特徴とする音声認識装置。

【考案の詳細な説明】

【0001】

【産業上の利用分野】 本考案は音声認識装置に関するものである。

【0002】

【従来の技術】 音声認識はパターン認識の一分野であって、音声の各音をパターン化して自動的手段によって識

3

別する。このため、音声入力部は或る程度の音声入力レベルの変化に追従できることが求められる。

【0003】従来、ある程度の音声入力レベルの変化に追従できるようにするため、音声入力により得た音声変換電圧信号を対数増幅器或いはコンパンダーなどによって圧縮処理して以後の信号処理部に送るようにしている。

【0004】

【考案が解決しようとする課題】上述したように従来の装置では、単に信号を圧縮しているだけであるため、音声-電気変換手段であるマイクロホンからの距離、声の大小によって音声変換電圧信号のレベルが著しく変化してしまい、以後の音声認識のための信号処理に大きな影響を及ぼし、認識率の低下を招くという問題点があった。

【0005】従って本考案は、上述した従来の問題点に鑑み、音声-電気変換手段からの音声変換電圧信号のレベルが該変換手段からの距離、声の大小によって変わっても、以後の音声認識部に一定レベルの信号を供給し認識率の向上を図った音声認識装置を提供することを目的としている。

【0006】

【課題を解決する手段】上記目的を達成するため本考案により成された音声認識装置は、音声入力のレベルに応じた大きさの音声変換電圧信号を発生する音声-電気変換手段と、該音声-電気変換手段からの音声変換電圧信号を増幅する増幅手段と、該増幅手段の利得を設定する利得設定手段と、前記増幅手段を介して前記音声-電気変換手段からの音声変換電圧信号を受け取り音声認識を行う音声認識部と、前記音声-電気変換手段からの音声変換電圧信号を予め定めた基準電圧と比較し、前記音声変換電圧信号が基準電圧以上となったときセット信号を出力する比較手段と、音声発声者の操作により発生されるリセット信号によりリセットされ、リセット状態にあるとき前記比較手段からのセット信号の入力により一定時間持続する時間信号を出力するタイマ手段と、該タイマ手段が時間信号を出力している間に入力する前記音声変換電圧信号のピーク値を検出するピーク検出手段とを備え、前記利得設定手段が、前記ピーク検出手段により検出したピーク値に反比例した値に前記増幅手段の利得を設定することを特徴としている。

【0007】

【作用】以上の構成において、音声を発生する前に音声発声者の操作により発生されるリセット信号でタイマ手段がリセットされる。リセットされたタイマ手段は、音声発声者の音声-電気変換手段からの音声変換電圧信号が予め定めた基準電圧以上となったとき比較手段が出力するセット信号の入力により一定時間持続する時間信号を出力し、この間に入力する音声変換電圧信号のピーク値をピーク検出手段が検出し、利得設定手段がピーク検

4

出手段により検出したピーク値に反比例した値に増幅手段の利得を設定する。以上のように、音声発声者が発声した音声の頭部分の音声ピークのレベルによってそれ以降の信号レベルを予想してこのピーク値に反比例した利得を増幅手段に設定しているので、増幅手段の利得は小さい音声のときには大きく、大きい音声のときには小さくされ、増幅手段により増幅されてその後続く音声に対応する音声変換電圧信号は常に一定レベルで音声認識部に供給されるようになり、音声-電気変換手段と発声者の距離や音声の大小によって誤認識が生じることが防止される。

【0008】

【実施例】以下、本考案による音声認識装置の一実施例を図面に基づいて説明する。図1は本考案による音声認識装置の一実施例を示すブロック図であり、11は入力端子であり、該端子11には図示しない音声-電気変換手段としてのマイクロホンによりピックアップされた音声に基づいて発生される音声変換電圧信号（以下音声信号と略記する）が入力される。

【0009】12は比較手段としてのコンパレータであり、該コンパレータ12の一方の入力には音声信号が、他方の入力には抵抗 R_1 、 R_2 によって分圧して得た基準電圧 V_R がそれぞれ印加される。13は例えばR-Sフリップフロップにより構成されるラッチ回路であり、そのセット入力にはコンパレータ12の出力からのセット信号が、リセット入力にはリセット入力端子13'からのリセット信号がそれぞれ入力される。14はラッチ回路13の出力によりトリガされ一定時間持続する時間信号を発生するワンショットタイマ回路、15はピークホールド回路であり、ワンショットタイマ回路14からの時間信号の期間に入力端子11からの音声信号のピーク値を検出し、そのレベルを保持して出力する。16はアナログ-デジタル(A-D)変換器であり、ピークホールド回路15のアナログ出力をデジタル化して出力する。17はデジタル化されたピーク値をデコードするデコーダ、18a~18dはデコーダ17の出力によりオン・オフ制御されるアナログスイッチであり、アナログスイッチ18a~18dの各々は抵抗 R_a ~ R_d が直列に接続された上で並列に接続されている。

【0010】19は増幅器であり、上記アナログスイッチ18a~18dと抵抗 R_a ~ R_d の並列回路を介して入力端子11に接続され、その出力が音声認識部20に接続されている。なお、 R_2 は増幅器19の帰還抵抗である。

【0011】上記並列回路はオンしているアナログスイッチ18a~18dの組合せにより、抵抗 R_a ~ R_d の並列回路を増幅器19の入力に直列に挿入し、増幅器19の入力信号レベル、ひいては増幅器19の利得を決定する。

【0012】以上の構成において、図2(a)に示すよ

5

うな音声信号 S_v が入力端子11に入力されたとする。入力端子11からの音声信号 S_v はコンパレータ12に印加され、ここで基準電圧 V_R と比較される。音声信号 S_v が時点 t_1 で基準電圧 V_R を越えると、コンパレータ12の出力にLからHレベルに立上るセット信号が発生され、このセット信号に応じてラッチ回路13にHレベルがラッチされ、その出力がLからHレベルに立上る。このラッチ回路13の出力の立上りにより、ワンショットタイマ回路14がトリガされ、その出力に図2(b)に示すように時点 t_2 までの一定時間 T の間Hレベルを持続する時間信号 S_T を出力し、これをピークホールド回路15に印加する。以上により、ラッチ回路13及びワンショットタイマ回路14は、音声発声者の操作により発生されるリセット信号によりリセットされ、リセット状態にあるときコンパレータ12からのセット信号の入力により一定時間持続する時間信号を出力するタイマ手段を構成している。

【0013】ピークホールド回路15は、時間信号 S_T が印加されている間音声信号 S_v のピーク値を検出し、その後そのピーク値を出力し続ける。図の例では時点 t_1 で検出したピーク値 V_p を出力する。ピークホールド回路15が出力するピーク値はA-D変換器16によりデジタル化された後、デコーダ17によりデコードされてアナログスイッチ18a~18dの制御信号に変換される。デコーダ17が発生する制御信号によりオンするアナログスイッチ18a~18dが決定され、抵抗 R_a ~ R_d の適宜のものが増幅器19の入力と入力端子11との間に挿入される。従って、増幅器19へは該挿入された抵抗を通じて音声信号 S_v が入力されることになる。

【0014】ところで、人間が発する音声は同じ単語であれば、各人に応じた独特の音声信号が常に得られる。従って、この音声信号の最初の部分のレベルを基準にしてその後の音声信号レベルを調整すれば、たとえ音声-電気変換手段と発声者の距離が変わったり、或いは声の大きさが変わっても、調整後の音声信号レベルは略一定にすることができ、この調整後の音声信号により音声認識を行えば誤認識はほとんど起らず、認識率は著しく向上する。

【0015】なお、リセット入力端子13'に入力されるリセット信号は、例えば発声者が発声に先だちスイッチを操作することによって発生することができ、このリセット信号の印加によりラッチ回路13がリセットされ

6

る毎に、その後の音声信号に応じて増幅器19の利得が制御されるようになる。

【0016】上述した本考案の装置によれば、図3(a)に示すように音声信号が小さいレベルのときも、図3(b)に示すように音声信号が大きいレベルのときも、増幅器19を通じて音声認識部20に入力される音声信号は図3(c)に示すような略同一のものとなる。従って、この図3(c)に示すような音声信号に対応する信号により予め基準パターンを登録しておくことにより、特定の人による音声入力があったか否かを判定し、例えば解錠信号などの制御信号を発生することができ、しかし、時点 t_1 から t_2 までの時間 T の間の信号は調整されておらず、互に異なるので、音声認識に当ってはこの部分は無視する。

【0017】

【考案の効果】以上説明したように本考案によれば、音声発声者が発声する前に行う操作によって発生されるリセット信号によりリセットされたタイマ手段をその後の発声によって発生されるセット信号によりセットして一定時間を発生させ、この一定時間の間に検出した音声のピーク値に反比例した利得を設定させるようにしているので、音声入力の大きさが違っていても、この設定した利得で増幅後に音声認識部に入力される音声変換電圧信号は常に一定の大きさとすることができるようになり、発声者との距離や音声の大小によって音声認識の結果が変わってしまうことがなく、音声認識が著しく向上するようになる。

【図面の簡単な説明】

【図1】図1は本考案による音声認識装置の一実施例を示すブロック図である。

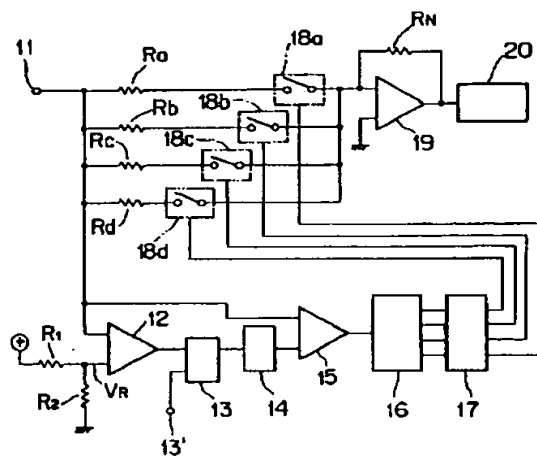
【図2】図2は図1の装置の各部の波形を示す図である。

【図3】図3は本考案の装置の動作を説明するための波形図である。

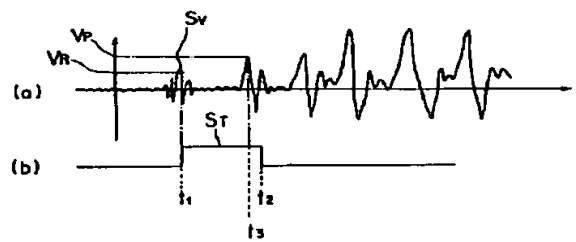
【符号の簡単な説明】

- 11 入力端子
- 12 コンパレータ
- 13 ラッチ回路
- 14 ワンショットタイマ回路
- 15 ピークホールド回路
- 16 A-D変換器
- 19 増幅器
- 20 音声認識部

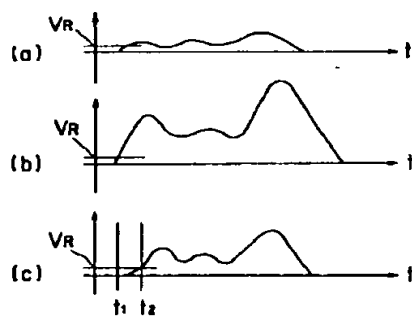
【図1】



【図2】



【図3】



Partial Translation of JP2500761U

[0005] In consideration of the aforementioned problem of the prior art, therefore, an object of the present invention is to provide a voice recognition unit supplying a signal of a constant level to a subsequent voice recognizer also when the level of a voice conversion voltage signal from voice-electric conversion means varies with the distance from the conversion means or the volume of the voice thereby improving the recognition rate.

[0006]

[Means for Solving the Problem] In order to attain the aforementioned object, a voice recognition unit according to the present invention comprises voice-electric conversion means generating a voice conversion voltage signal having a magnitude responsive to the level of a voice input, amplification means amplifying the voice conversion voltage signal from the voice-electric conversion means, gain setting means setting the gain of the amplification means, a voice recognizer receiving the voice conversion voltage signal from the said voice-electric conversion means through the said amplification means and performing voice recognition, compare means comparing the voice conversion voltage signal from the said voice-electric conversion means with a predetermined reference voltage and outputting a set signal when the said

voice conversion voltage signal exceeds the reference voltage, timer means reset by a reset signal generated by manipulation of a speaker for outputting a time signal sustained for a constant time by receiving the set signal from the said compare means in the reset state and peak detection means detecting the peak of the said voice conversion voltage signal received while the timer means outputs the time signal, and the said gain set means sets gain of the said amplification means to a value inversely proportional to the peak detected by the said peak detection means.

[0007]

[Function] In the aforementioned structure, the timer means is reset by the reset signal generated by manipulation of the speaker before speaking. The reset timer means outputs the time signal sustained for the constant time by receiving the set signal output from the compare means when the voice conversion voltage signal from the voice-electric conversion means responsive to the speaker exceeds the predetermined reference voltage, the peak detection means detects the peak of the currently received voice conversion voltage signal, and the gain set means sets the gain of the amplification means to the value inversely proportional to the peak detected by the peak detection means. As hereinabove described, the subsequent signal level is predicted in response to the level of the peak of the head of the speaker's voice for setting the

gain inversely proportional to this peak in the amplification means, whereby the gain of the amplification means is rendered large for a small voice and rendered small for a large voice and the voice conversion voltage signal amplified by the amplification means to correspond to the subsequent voice is regularly supplied to the voice recognizer at a constant level, for preventing false recognition resulting from the distance between the voice-electric conversion means and the speaker or the volume of the voice.

10 [0008]

[Embodiment] An embodiment of the voice recognition unit according to the present invention is now described with reference to the drawings. Fig. 1 is a block diagram showing the embodiment of the voice recognition unit according to the present invention, and numeral 11 denotes an input terminal receiving a voice conversion voltage signal (hereinafter simply referred to as a voice signal) generated on the basis of a voice picked up by a microphone (not shown) serving as voice-electric conversion means.

20 [0009] Numeral 12 denotes a comparator serving as compare means, and the voice signal and a reference voltage V_R obtained by performing division with resistors R_1 and R_2 are applied to first and second inputs of the comparator 12 respectively. Numeral 13 denotes a latch circuit formed by an R-S flip-flop, for example, and a set signal from an output of the comparator 12

25

and a reset signal from a reset input terminal 13' are input in a set input and a reset input thereof respectively. Numeral 14 denotes a one-shot time circuit triggered by an output of the latch circuit 13 for generating a time signal sustained for a constant time, and numeral 15 denotes a peak hold circuit detecting the peak of the voice signal from the input terminal 11 in the period of the time signal from the one-shot timer circuit 14, and holding the level thereof and outputting the same. Numeral 16 denotes an analog-digital (A-D) converter digitizing an analog output from the peak hold circuit 15 and outputting the same. Numeral 17 denotes a decoder decoding the digitized peak and numerals 18a to 18d denote analog switches on-off controlled by an output from the decoder 17, while the analog switches 18a to 18d are connected in parallel with serially connected resistors Ra to Rd.

[0010] Numeral 19 denotes an amplifier connected to the input terminal 11 through the aforementioned parallel circuits of the analog switches 18a to 18d and the resistors Ra to Rd, so that its output is connected to the voice recognizer 20. Symbol R_N denotes feedback resistance of the amplifier 19.

[0011] The parallel circuits of the resistors Ra to Rd are serially inserted in the inputs of the amplifier 19 by combination of on-state analog switches 18a to 18d, for deciding the input signal level of the amplifier 19 as well as the gain of the amplifier 19.

[0012] It is assumed that a voice signal S_v shown in Fig. 2(a) is input in the input terminal 11 in the aforementioned structure. The voice signal S_v from the input terminal 11 is applied to the comparator 12, and compared with the reference voltage V_R . When the voice signal S_v exceeds the reference voltage V_R at a point t_1 , a set signal going high from low is generated at the output of the comparator 12, and the latch circuit 13 latches a high level in response to this set signal so that its output goes high from low. The one-shot timer circuit 14 is triggered due to this rising of the output of the latch circuit 13 for outputting a time signal S_t sustaining a high level for a constant time up to a point t_2 as shown in Fig. 2(b) on its output and applying the same to the peak hold circuit 15. Thus, the latch circuit 13 and the one-shot timer circuit 14 form timer means reset by the reset signal generated by manipulation of the speaker for outputting the time signal sustained for a constant time by receiving the set signal from the comparator 12 in the reset state.

[0013] The peak hold circuit 15 detects the peak of the voice signal S_v while the time signal S_t is applied, and thereafter continuously outputs the peak. In the illustrated example, the peak hold circuit 15 outputs a peak V_p detected at the point t_2 . The peak output from the peak hold circuit 15 is digitized by the A-D converter 16 and thereafter decoded by the decoder 17 to be converted to a control signal for the analog switches

18a to 18d. One of the analog switches 18a to 18d to be turned on is decided by the control signal generated by the decoder 17, and a proper one of the resistors Ra to Rd is inserted between the input of the amplifier 19 and the input terminal 11.

5 Therefore, the voice signal S_v is input in the amplifier 19 through the inserted resistor.

[0014] When a human speaks the same word, a specific voice signal responsive to each speaker is regularly obtained. When a subsequent voice signal level is adjusted with reference to
10 the level of the initial part of this voice signal, therefore, the adjusted voice signal level can be rendered substantially constant even if the distance between the voice-electric conversion means and the speaker or the volume of the voice varies, while false recognition hardly results and the
15 recognition rate is remarkably improved when voice recognition is performed with the adjusted voice signal.

[0015] The speaker can generate the reset signal input in the reset input terminal 13' by manipulating a switch before speaking, for example, so that the gain of the amplifier 19
20 is controlled in response to the subsequent voice signal every time the latch circuit 13 is reset by application of this reset signal.

[0016] According to the aforementioned unit according to the present invention, a substantially identical voice signal is
25 input in the voice recognizer 20 through the amplifier 19 as

shown in Fig. 3(c) whether the voice signal is at a small level as shown in Fig. 3(a) or the voice signal is at a large level as shown in Fig. 3(b). When a reference pattern is previously registered by a signal corresponding to the voice signal shown in Fig. 3(c), therefore, it is possible to determine whether or not a voice is input by a specific person for generating a control signal such as a release signal, for example. However, signals in a time T between the points t_1 and t_2 are not adjusted but different from each other, and hence this part is ignored in voice recognition.

[0017]

[Effect of the Invention] As hereinabove described, the unit according to the present invention sets the timer means reset by the reset signal by the manipulation made by the speaker before speaking with the set signal generated by subsequent speaking for generating a constant time and sets the gain inversely proportional to the peak of the voice detected in the constant time, whereby the voice conversion voltage signal input in the voice recognizer after amplification with the set gain can be regularly set to a constant level so that the result of voice recognition does not vary with the distance to the speaker or the volume of the voice but voice recognition is remarkably improved.

(19)日本国特許庁 (J P)

(12) 実用新案登録公報 (Y 2) (11)実用新案登録番号

第2500761号

(45)発行日 平成 8 年(1996) 6 月12日

(24)登録日 平成 8 年(1996) 3 月28日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/02	3 0 1		G 1 0 L 3/02	3 0 1 B

(全 4 頁)

(21)出願番号 実願平6-3041
特願昭61-122324の変更
(22)出願日 昭和61年(1986) 5 月29日
(65)公開番号 実開平6-82700
(43)公開日 平成 6 年(1994)11月25日

(73)実用新案権者 000170598
株式会社アルファ
神奈川県横浜市金沢区福浦 1 丁目 6 番 8
号
(72)考案者 臼井 俊之
東京都大田区蒲田 2 丁目 8 番 2 号 国産
金属工業株式会社内
(74)代理人 弁理士 瀧野 秀雄 (外 1 名)

審査官 渡邊 聡

(56)参考文献 特開 昭57-90750 (J P, A)
特開 昭58-70287 (J P, A)
特開 昭57-179898 (J P, A)
特開 昭58-125099 (J P, A)

(54)【考案の名称】 音声認識装置

1

(57)【実用新案登録請求の範囲】

【請求項 1】 音声入力レベルに応じた大きさの音声
変換電圧信号を発生する音声-電気変換手段と、
該音声-電気変換手段からの音声変換電圧信号を増幅す
る増幅手段と、

該増幅手段の利得を設定する利得設定手段と、

前記増幅手段を介して前記音声-電気変換手段からの音
声変換電圧信号を受け取り音声認識を行う音声認識部
と、

前記音声-電気変換手段からの音声変換電圧信号を予め
10 定めた基準電圧と比較し、前記音声変換電圧信号が基準
電圧以上となったときセット信号を出力する比較手段
と、

音声発声者の操作により発生されるリセット信号により
リセットされ、リセット状態にあるとき前記比較手段か

2

らのセット信号の入力により一定時間持続する時間信号
を出力するタイマ手段と、

該タイマ手段が時間信号を出力している間に入力する前
記音声変換電圧信号のピーク値を検出するピーク検出手
段とを備え、

前記利得設定手段が、前記ピーク検出手段により検出し
たピーク値に反比例した値に前記増幅手段の利得を設定
することを特徴とする音声認識装置。

【考案の詳細な説明】

【0001】

【産業上の利用分野】 本考案は音声認識装置に関するも
のである。

【0002】

【従来の技術】 音声認識はパターン認識の一分野であっ
て、音声の各音をパターン化して自動的手段によって識

別する。このため、音声入力部は或る程度の音声入力レベルの変化に追従できることが求められる。

【0003】従来、ある程度の音声入力レベルの変化に追従できるようにするため、音声入力により得た音声変換電圧信号を対数増幅器或いはコンパンダーなどによって圧縮処理して以後の信号処理部に送るようにしている。

【0004】

【考案が解決しようとする課題】上述したように従来の装置では、単に信号を圧縮しているだけであるため、音声-電気変換手段であるマイクロホンからの距離、声の大小によって音声変換電圧信号のレベルが著しく変化してしまい、以後の音声認識のための信号処理に大きな影響を及ぼし、認識率の低下を招くという問題点があった。

【0005】従って本考案は、上述した従来の問題点に鑑み、音声-電気変換手段からの音声変換電圧信号のレベルが該変換手段からの距離、声の大小によって変わっても、以後の音声認識部に一定レベルの信号を供給し認識率の向上を図った音声認識装置を提供することを目的としている。

【0006】

【課題を解決する手段】上記目的を達成するため本考案により成された音声認識装置は、音声入力のレベルに応じた大きさの音声変換電圧信号を発生する音声-電気変換手段と、該音声-電気変換手段からの音声変換電圧信号を増幅する増幅手段と、該増幅手段の利得を設定する利得設定手段と、前記増幅手段を介して前記音声-電気変換手段からの音声変換電圧信号を受け取り音声認識を行う音声認識部と、前記音声-電気変換手段からの音声変換電圧信号を予め定めた基準電圧と比較し、前記音声変換電圧信号が基準電圧以上となったときセット信号を出力する比較手段と、音声発声者の操作により発生されるリセット信号によりリセットされ、リセット状態にあるとき前記比較手段からのセット信号の入力により一定時間持続する時間信号を出力するタイマ手段と、該タイマ手段が時間信号を出力している間に入力する前記音声変換電圧信号のピーク値を検出するピーク検出手段とを備え、前記利得設定手段が、前記ピーク検出手段により検出したピーク値に反比例した値に前記増幅手段の利得を設定することを特徴としている。

【0007】

【作用】以上の構成において、音声が発生する前に音声発声者の操作により発生されるリセット信号でタイマ手段がリセットされる。リセットされたタイマ手段は、音声発声者の音声-電気変換手段からの音声変換電圧信号が予め定めた基準電圧以上となったとき比較手段が出力するセット信号の入力により一定時間持続する時間信号を出力し、この間に入力する音声変換電圧信号のピーク値をピーク検出手段が検出し、利得設定手段がピーク検

出手段により検出したピーク値に反比例した値に増幅手段の利得を設定する。以上のように、音声発声者が発声した音声の頭部分の音声ピークのレベルによってそれ以降の信号レベルを予想してこのピーク値に反比例した利得を増幅手段に設定しているので、増幅手段の利得は小さい音声のときには大きく、大きい音声のときには小さくされ、増幅手段により増幅されてその後続く音声に対応する音声変換電圧信号は常に一定レベルで音声認識部に供給されるようになり、音声-電気変換手段と発声者の距離や音声の大小によって誤認識が生じることが防止される。

【0008】

【実施例】以下、本考案による音声認識装置の一実施例を図面に基づいて説明する。図1は本考案による音声認識装置の一実施例を示すブロック図であり、11は入力端子であり、該端子11には図示しない音声-電気変換手段としてのマイクロホンによりピックアップされた音声に基づいて発生される音声変換電圧信号（以下音声信号と略記する）が入力される。

【0009】12は比較手段としてのコンパレータであり、該コンパレータ12の一方の入力には音声信号が、他方の入力には抵抗 R_1 、 R_2 によって分圧して得た基準電圧 V_R がそれぞれ印加される。13は例えばR-Sフリップフロップにより構成されるラッチ回路であり、そのセット入力にはコンパレータ12の出力からのセット信号が、リセット入力にはリセット入力端子13'からのリセット信号がそれぞれ入力される。14はラッチ回路13の出力によりトリガされ一定時間持続する時間信号を発生するワンショットタイマ回路、15はピークホールド回路であり、ワンショットタイマ回路14からの時間信号の期間に入力端子11からの音声信号のピーク値を検出し、そのレベルを保持して出力する。16はアナログ-デジタル（A-D）変換器であり、ピークホールド回路15のアナログ出力をデジタル化して出力する。17はデジタル化されたピーク値をデコードするデコーダ、18a~18dはデコーダ17の出力によりオン・オフ制御されるアナログスイッチであり、アナログスイッチ18a~18dの各々は抵抗 R_a ~ R_d が直列に接続された上で並列に接続されている。

【0010】19は増幅器であり、上記アナログスイッチ18a~18dと抵抗 R_a ~ R_d の並列回路を介して入力端子11に接続され、その出力が音声認識部20に接続されている。なお、 R_n は増幅器19の帰還抵抗である。

【0011】上記並列回路はオンしているアナログスイッチ18a~18dの組合せにより、抵抗 R_a ~ R_d の並列回路を増幅器19の入力に直列に挿入し、増幅器19の入力信号レベル、ひいては増幅器19の利得を決定する。

【0012】以上の構成において、図2（a）に示すよ

5

うな音声信号 S_v が入力端子11に入力されたとする。入力端子11からの音声信号 S_v はコンパレータ12に印加され、ここで基準電圧 V_R と比較される。音声信号 S_v が時点 t_1 で基準電圧 V_R を越えると、コンパレータ12の出力にLからHレベルに立上るセット信号が発生され、このセット信号に応じてラッチ回路13にHレベルがラッチされ、その出力がLからHレベルに立上る。このラッチ回路13の出力の立上りにより、ワンショットタイマ回路14がトリガされ、その出力に図2(b)に示すように時点 t_2 までの一定時間Tの間Hレベルを持続する時間信号 S_T を出力し、これをピークホールド回路15に印加する。以上により、ラッチ回路13及びワンショットタイマ回路14は、音声発声者の操作により発生されるリセット信号によりリセットされ、リセット状態にあるときコンパレータ12からのセット信号の入力により一定時間持続する時間信号を出力するタイマ手段を構成している。

【0013】ピークホールド回路15は、時間信号 S_T が印加されている間音声信号 S_v のピーク値を検出し、その後そのピーク値を出力し続ける。図の例では時点 t_1 で検出したピーク値 V_1 を出力する。ピークホールド回路15が出力するピーク値はA-D変換器16によりデジタル化された後、デコーダ17によりデコードされてアナログスイッチ18a~18dの制御信号に変換される。デコーダ17が発生する制御信号によりオンするアナログスイッチ18a~18dが決定され、抵抗 R_a ~ R_d の適宜のものが増幅器19の入力と入力端子11との間に挿入される。従って、増幅器19へは該挿入された抵抗を通じて音声信号 S_v が入力されることになる。

【0014】ところで、人間が発する音声は同じ単語であれば、各人に応じた独特の音声信号が常に得られる。従って、この音声信号の最初の部分のレベルを基準にしてその後の音声信号レベルを調整すれば、たとえ音声-電気変換手段と発声者の距離が変わったり、或いは声の大きさが変わっても、調整後の音声信号レベルは略一定にすることができ、この調整後の音声信号により音声認識を行えば誤認識はほとんど起らず、認識率は著しく向上する。

【0015】なお、リセット入力端子13'に入力されるリセット信号は、例えば発声者が発声に先だちスイッチを操作することによって発生することができ、このリセット信号の印加によりラッチ回路13がリセットされ

6

る毎に、その後の音声信号に応じて増幅器19の利得が制御されるようになる。

【0016】上述した本考案の装置によれば、図3(a)に示すように音声信号が小さいレベルのときも、図3(b)に示すように音声信号が大きいレベルのときも、増幅器19を通じて音声認識部20に入力される音声信号は図3(c)に示すような略同一のものとなる。従って、この図3(c)に示すような音声信号に対応する信号により予め基準パターンを登録しておくことにより、特定の人による音声入力があったか否かを判定し、例えば解錠信号などの制御信号を発生することができ、しかし、時点 t_1 から t_2 までの時間Tの間の信号は調整されておらず、互に異なるので、音声認識に当ってはこの部分は無視する。

【0017】

【考案の効果】以上説明したように本考案によれば、音声発声者が発声する前に行う操作によって発生されるリセット信号によりリセットされたタイマ手段をその後の発声によって発生されるセット信号によりリセットして一定時間を発生させ、この一定時間の間に検出した音声のピーク値に反比例した利得を設定させるようにしている。従って、音声入力の大きさが違っていても、この設定した利得で増幅後に音声認識部に入力される音声変換電圧信号は常に一定の大きさとするようにでき、発声者との距離や音声の大きさによって音声認識の結果が変わってしまうことがなく、音声認識が著しく向上するようになる。

【図面の簡単な説明】

【図1】図1は本考案による音声認識装置の一実施例を示すブロック図である。

【図2】図2は図1の装置の各部の波形を示す図である。

【図3】図3は本考案の装置の動作を説明するための波形図である。

【符号の簡単な説明】

- 11 入力端子
- 12 コンパレータ
- 13 ラッチ回路
- 14 ワンショットタイマ回路
- 15 ピークホールド回路
- 16 A-D変換器
- 19 増幅器
- 20 音声認識部

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-115098

(43)Date of publication of application : 07.05.1996

(51)Int.Cl. G10L 9/00
G10L 5/02

(21)Application number : 06-251124

(71)Applicant : HITACHI MICROCOMPUT SYST LTD

(22)Date of filing : 18.10.1994

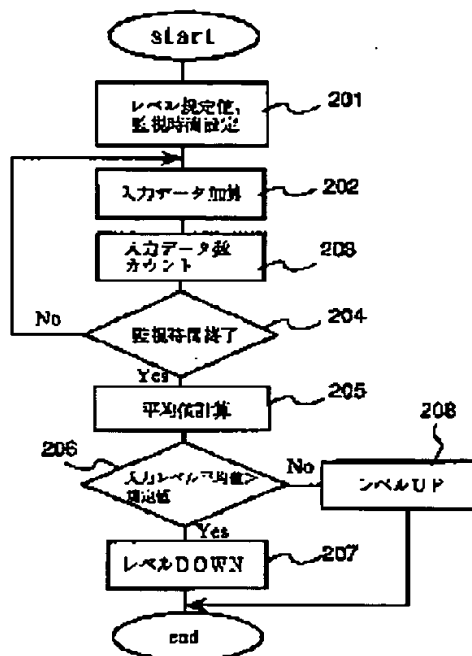
(72)Inventor : MORITA HIROSHI
TSUNODA TOMOAKI
MIURA AKIYOSHI
AOKI HISANOBU

(54) METHOD AND DEVICE FOR EDITING VOICE

(57)Abstract:

PURPOSE: To provide a voice editing technique capable of monitoring an output value of a Codec previous to recording, setting an input level in an optimum value by a comparison result with a prescribed value and reducing a recorded data amount in accordance with the variation of the number of quantization bits of the sampling data in a system coding a voice and storing the data in a computer.

CONSTITUTION: This device is a voice editing device constituted of a CPU, a memory, a voice source, an A/D converter, a Codec interface, a disk interface, a disk, a speaker and a D/A converter, etc. After a level prescribed value and a monitoring time are set first of all, the digital data are fetched, and the addition of the data and the counting of the number of data are performed for a fixed time (201-204). Then, a mean of input values is calculated after the prescribed time is over, and the mean value is compared with the prescribed value, and when the mean value is larger, an input level is set in lower, and on the other hand, when the mean value is smaller, the input level is set in higher.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]